

## (12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau(43) International Publication Date  
20 December 2001 (20.12.2001)

PCT

(10) International Publication Number  
**WO 01/97211 A2**(51) International Patent Classification<sup>7</sup>: **G10L**

(21) International Application Number: PCT/US01/41031

(22) International Filing Date: 18 June 2001 (18.06.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/212,319 16 June 2000 (16.06.2000) US(71) Applicant (for all designated States except US): **HEALTHETECH, INC.** [US/US]; 523 Park Point Dr., 3rd Floor, Golden, CO 80401 (US).

(72) Inventor; and

(75) Inventor/Applicant (for US only): **MAULT, James, R.** [US/US]; 1580 Blakcomb Court, Evergreen, CO 80439 (US).(74) Agents: **KRASS, Allen, M.** et al.; Gifford, Krass, Groh, Sprinkle, Anderson & Citkowski, PC, Suite 400, 280 N. Old Woodward Ave., Birmingham, MI 48009 (US).

(81) Designated States (national): AE, AG, AI, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

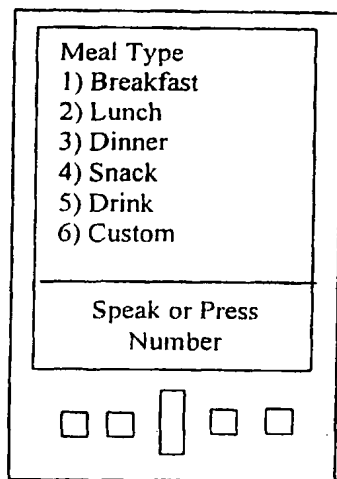
(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: SPEECH RECOGNITION CAPABILITY FOR A PERSONAL DIGITAL ASSISTANT



(57) Abstract: A speech recognition module for a personal digital assistant comprises: a module housing designed to engage with an accessory feature of the PDA, such as an accessory slot; a microphone for receiving speech commands from the person; and a speech recognition system. A corresponding electrical speech command signal is communicated to the portable computing device, allowing control of the operation of a software application program running on the portable computing device. In particular, menu items can be selected for generation of a diet log for the person, for example during a weight control program.

WO 01/97211 A2

# SPEECH RECOGNITION CAPABILITY FOR A PERSONAL DIGITAL ASSISTANT

## Field of the Invention

The invention relates to speech recognition capabilities of a computing device. In particular, the invention relates to the use of speech recognition in creating a diet log for a person on a weight control program.

## Background of the Invention

Portable computing devices such as personal digital assistants (PDAs), including Palm Pilots, Handspring Visors, hand-held PCs, and similar devices, have become incredibly popular because of their functionality and flexibility. For purposes of this application, the term computing device should be understood to include all computing devices, including those with additional functionality, including: desk-top, laptop, hand-held, palm top, and wearable computers; and also cellular phones, telephones, two-way paging devices, radios, communicators, calculators, organizers, e-books, navigation devices, automobile electronics, digital interactive television controllers, and other electronic devices having computing capability. However, for ease of description, this application will focus on portable computing devices such as the Palm Pilot family of hand-held devices (Palm Inc., Santa Clara, CA) and the Handspring Visor (Handspring, Mountain View, CA).

These personal digital assistants (or PDAs) include buttons and touch sensitive screens allowing input of data in a variety of ways, including various forms of handwriting recognition and simplified keyboards. For some persons and applications, these forms of data input are insufficient. For example, a person who wishes to input information, such as recording what meals were eaten during a day or expenses incurred on a trip, must grasp an input stylus with one hand while holding the PDA in the other hand. The stylus is then used to touch the touch sensitive screen in a manner indicative of the input desired. This obviously ties up both of the person's hands, and requires a moderate level of visual and mental focus. It would be more desirable if data input could be made using a person's voice, rather than manual manipulation.

- 2 -

A partial solution is offered by some hand-helds which include voice recording capability. A person who wishes to quickly record information on the PDA presses a record button and speaks into a microphone. Later, the recorded information may be played back. Unfortunately, this approach falls far short of what is desirable. The voice recording is only voice-to-voice and therefore someone must later listen to the recorded voice and manually input the data. While the original recording step may be easier than manual manipulation, manual manipulation is not avoided because it must be completed at a later time. More desirable would be a voice-to-text voice recognition capability so that a person may speak commands and information and the PDA responds to the commands and records the information in text form. However, voice-to-text speech recognition is extremely complex and requires powerful computer processors and a large data storage capability. Current PDAs offer only a fraction of the necessary processing power and storage.

Speech recognition software and systems are well known in the art, for example as described in U.S. Pat. Nos. 5,749,072 to Mazurkiewicz et al., and 6,167,255 to Kennedy III et al, incorporated herein by reference.

### Summary of the Invention

According to the present invention, the limitations of currently available computing devices (such as portable computing devices including PDAs) and speech recognition software are overcome by offering a simplified speech recognition program and software designed to work with this simplified speech recognition capability. Much of the complexity of speech recognition software comes from the huge variety of words in a given language and the similarity in sounds of many of these words. The software must be capable of very accurately capturing the spoken sounds. Then, the software must determine what word was spoken out of a variety of similar choices by analyzing the context of the word, statistical likelihood, and past usage by a given person. Also, persons do not consistently pronounce words so software must be "trained" to recognize a specific person. The speech recognition challenge is significantly reduced by artificially constraining the vocabulary the software and hardware must interpret. Examples of restricted vocabulary speech recognition are currently used on many telephone systems, allowing persons to either

- 3 -

press a number 0 through 9 or to speak the number. In this situation, the speech recognition software must only recognize and distinguish between ten choices, all of which are reasonably distinct.

According to the present invention, a similar capability is provided to PDAs.

5 In a preferred embodiment, the PDA is capable of recognizing numbers 0 through 9 when they are spoken by a person. Current PDAs do not have this hardware or software capability. However, software may be added to most PDAs. Also, most PDAs are capable of accepting add-on hardware modules. For example, the Handspring Visor has a slot arrangement on its back side known as a Springboard. A

10 variety of hardware modules, such as memory modules, may be inserted into the Springboard to enhance the capability of the Handspring Visor. Likewise, the Palm Pilot family of hand-held PDAs accept clip-on hardware modules to enhance their capabilities. It is expected that future versions of PDAs will include new and simplified hardware attachment capabilities.

15 Typically, a hardware module which is inserted into or attached to a PDA may also include application software further enhancing the capabilities of the PDA and taking advantage of the additional hardware. For example, the Handspring Visor accepts numerous Springboard compatible hardware modules, such as digital camera modules and music recording and playback modules. These modules interconnect

20 with the Springboard portion of the Handspring Visor and include hardware such as a digital camera lens and related electronics, or music recording and playback hardware. This add-on hardware cooperates with the existing capabilities of the Handspring Visor to provide the capabilities of a digital camera or digital music player and recorder. The hardware modules also include onboard software which is

25 accessed and used by the hardware of the Handspring Visor and the add-on hardware module. The software allows the module and Handspring Visor to communicate and provide the enhanced capabilities. Likewise, hardware modules such as wireless modems may be interconnected with some members of the Palm Pilot family.

A speech recognition module for a personal digital assistant comprises: a

30 module housing configured to engage with an accessory feature of the PDA; a speech receiving device for receiving speech commands from the person; and a speech recognition means capable of distinguishing multiple speech commands and

- 4 -

communicating a corresponding set of electrical commands to the PDA. Further, the speech recognition module for a portable computing device can comprise: a module housing configured to engage with an accessory interface of the portable computing device; a speech receiving device, adapted to receive a speech command from the person and to provide an electrical speech signal; and a speech recognition system, adapted to receive the electrical speech signal from the speech receiving device, to distinguish the speech command, and to communicate a corresponding electrical speech command signal to the portable computing device, wherein the electrical speech command signal is used to control the operation of a software application program running on the portable computing device. The speech receiving device can be a microphone. Speech recognition of microphone signals is disclosed in U.S. Pat. No. 5,749,072 to Mazurkiewicz et al., for example with reference to Figure 3 of that patent. The module can further comprise a nutritional database of food item identifiers and associated nutritional data for a number of food items. Nutritional data can comprise calorie content, fat content, fiber content, protein content, glycemic index, mineral content, vitamin content, and other nutritional components associated with each of the food item identifiers. The software application program can be adapted to select a selected food item identifier from the nutritional database under control of the electrical speech command signal, and to store the selected food item identifier and associated nutritional data in a memory so as to create a diet log for the person.

A portable computing system according to the present invention comprises: a display; a microprocessor; a memory; a database of nutritional data relating to each of a plurality of food item identifiers; a speech recognition system providing an electrical speech command signal correlated with a spoken input; and a software application program, executed by the microprocessor, adapted to display a food menu of food item identifiers on the display, to select a food item identifier from the food menu based on the electrical speech command signal received from the speech recognition system, and to store nutritional data relating to the selected food item identifier in the memory so as to create a diet log of selected food items. The nutritional database can be arranged in a hierarchical form in that particular food items are characterized by a generic type, and in the software application program can be further adapted to

- 5 -

display a menu of generic types on the display, and to select a generic type from the menu of generic types on receipt of the electrical speech signal.

An improved accessory module for a portable computing device (the combined system which can then be used for diet logging as part of a weight control  
5 program) comprises: an electrical interface, adapted to communicate with an accessory port of the portable computing device; a memory; a nutritional database, stored in the memory, correlating each of a number of food item identifiers with an associated food item nutritional content; and a software application program, stored in the memory, and executable by the microprocessor so as to present a menu of food  
10 item identifiers on the display, further adapted to select a food item selection from the menu of food item identifiers on receipt of a speech command signal, and further adapted to store the food item selection and correlated nutritional content within a diet log for the person. Nutritional content may included calorie content, diet component content, and other nutritional component of the food item. The speech command  
15 signal can be provided by a speech recognition system resident on the accessory module, on the portable computing device, or on another device in communication with the portable computing device.

A method of creating a diet log for a person, comprises: presenting a menu of food type identifiers to the person; receiving a type selection speech command from  
20 the person; correlating the type selection speech command with a selected food type identifier; presenting a menu of food item identifiers to the person, wherein the food item identifiers have a hierarchical relationship to the selected food type identifier; receiving an item selection speech command from the person; correlating the item selection speech command with a selected food item identifier; and storing the  
25 selected food item identifier and associated nutritional data within a memory device, so as to create a diet log for the person. The selected food item identifier is correlated with associated nutritional data using a nutritional database. The menus can be presented to the person on the display of a portable electronic device carried by the person, such as a personal digital assistant (PDA), other portable computing device,  
30 wrist-mounted device, wearable computer system, electronic book, electronic notebook, digital organizer, wireless phone, desktop computer system, web-TV, digital interactive TV, tablet computer, visor-mounted display system, or the like. The

- 6 -

menus presented to the person can be chosen based on the physical location of the person, for example as input by the person, as transmitted from a local wireless network, or as determined by a position location system such as a global positioning system (GPS). The correlation of speech commands with selected identifiers can  
5 performed by a software application program running on the portable computing device, accessory module to the portable computing device, or other device in communication with the portable computing device such as a remote computer system in communication with the portable computing device over a communications network.

10 In other embodiments, spoken commands can be recorded in a memory and analyzed later.

Computing devices may further include systems such as a separate speech receiving device (for example, a wrist mounted device), processor module (for example a belt mounted computer comprising a processor, memory, and plug-in  
15 hardware module comprising a speech recognition system), and display device (for example, a visor-mounted display). Preferably components are in wireless communication and function as a unitary device.

Wireless transmission methods can include Bluetooth, a wireless protocol developed by Ericsson (Sweden) and others, IEEE802.11(b) and similar, local  
20 wireless networks such as wireless Ethernet, IR, optical, and ultrasound methods. Wireless transmission methods between devices offers improved convenience, however cable connections can be used in place of wireless methods.

U.S. provisional application Serial No. 60/212,319 to James R. Mault (filed 6/16/00) is incorporated herein by reference.

## 25 Brief Description of the Drawings

Figure 1 shows a hardware module being inserted into a portable computing device;

Figure 2 shows a portable computing device having an inserted hardware module;

30 Figure 3 shows a portable computing device having an inserted hardware module;

- 7 -

Figures 4-6 show possible menu presentations on the display of a device, by which a person can select a cereal name to add to a diet log;

Figures 7-9 show other possible menu presentations on the display of a device, by which a person can select a cereal name to add to a diet log;

5        Figure 10A shows a simplified schematic of a portable computing device which may be used in embodiments of the present invention;

Figure 10B shows a simplified schematic of a hardware module which may be used in embodiments of the present invention;

Figure 11 is a flowchart illustrating a method of diet logging;

10       Figure 12A shows a presentation on a display, whereby a recognized letter is used to restrict menu presentation;

Figure 12B shows a further presentation on a display, whereby a recognized letter is used to restrict menu presentation; and

15       Figure 13 shown a schematic of a system in which the nutritional database and speech recognition system are accessed using a communications link to a remote server.

#### Detailed Description of the Invention

According to the present invention, a hardware module is provided which interconnects with a portable computing device, such as a PDA. The hardware  
20       module includes a microphone for receiving speech and may include other controls such as an on-off switch and sensitivity adjustments. An example of the present invention is shown schematically in Figures 1 and 2. The portable computing device 10 receives a hardware module 12 which, in the Figures, slides into a slot 16 in the back side of the portable computing device 10. The hardware module 12 includes a  
25       microphone 14 located near the edge of the module so that it is on the upper side of the portable computing device in use. In portable computing devices already including a microphone, the microphone may be omitted from the module. The module preferably includes both hardware and software so that when it is inserted in the slot in the portable computing device, the portable computing device becomes  
30       capable of simplified speech recognition.



- 8 -

Figure 3 shows another embodiment, in which module 26 is inserted into a slot in the lower edge of a portable computing device 20. Part of the module is exposed, allowing a person to speak into microphone 28. The portable computing device has display 22 and data entry keys 24. The display is divided into a menu display area 22a and a user instruction area 22b. However, this division is not critical to embodiments of the present invention.

Referring now to Figures 4-6, an example use for the simplified speech recognition module will be explained. In the Figures, a portable computing device is shown displaying textual menus for providing a number of choices. The hardware module is not shown. In this example, the portable computing device is being used to record dietary intake. In Figure 4, the portable computing device presents a menu with six options for meal type. Software providing the dietary intake software may be part of the hardware module, an additional module, or may be software loaded into, or already part of, the portable computing device. Using the speech recognition module, the person merely speaks the number of their choice of meal type. For example, if the person wishes to input information related to breakfast, they say the word "one." This is equivalent to entering the number 1 using the stylus or button on the portable computing device. This choice brings up a second menu, as shown in Figure 5. In this menu, several breakfast choices are displayed. If the person wishes to choose hot cereal, they pronounce the word "two." This, once again, enters the choice into the software causing the software to bring up a third menu, as shown in Figure 6. This third menu presents a variety of choices for hot cereal. It will be clear to those of skill in the art how this process may continue, in a very simplified manner, allowing a person to choose items for their diet. The chosen items, and associated nutritional data, can then be stored in a diet log for the person. This method may be similarly applied to data entry chores such as expense account tracking.

As known to those of skill in the art, a very wide variety of software applications are available for portable computing devices. According to the present invention, the hardware module and the software it carries are designed to work with many of the software applications. That is, the hardware and software module integrates with and communicates with a portable computing device such that speaking a number provides the same input to the software running on the portable

- 9 -

computing device as if the person had used a stylus to input the number. Also, additional vocabulary may be provided in the same hardware module, or another hardware module according to the present invention. For example, many portable computing device functions require the use of up and down buttons. For this purpose, the hardware module may also recognize the commands "up" and "down.". Commands could also be recognized for a few of the most commonly used functions of the portable computing device, such as "date" for date book, "to do" for the to do list, and "address" or "phone" for the address list, and to change operating modes of the portable computing device. The hardware module may also recognize a few letters such as A through J, or the entire alphabet, a set of alphanumeric characters, or a set of words corresponding to characters, for example alpha, bravo etc. for the alphabet A-Z. In addition, a small number of extra words such as back, delete, enter, and the like, for simple navigation can be added to the limited list recognized. Obviously, the speech recognition capabilities are entirely dependent on the complexity of the hardware and software in the module. Therefore, for some applications a very simple speech recognition module may be sufficient. In this case, recognizing numerals 0 through 9 may be more than sufficient. As the software and hardware capabilities of portable computing devices improve, more complex modules may be desirable for some applications.

The hardware module can be a proprietary accessory to a portable computing device, or may comply with well known standards, for example PCMCIA and Miniature Card standards. The design of an accessory module with imaging capabilities, which may be advantageously modified for use in the present invention by inclusion of a voice recognition system, is disclosed in U.S. Pat. No. 6,219,560 to Erkkila et al., incorporated herein by reference.

As will be clear to those of skill in the art, the present invention provides numerous advantages in the input of data. In the example given earlier in entering diet information, the person may simply enter their choices by quickly repeating the words "one," "two," "one." This may be both easier and more pleasant than manual input for many persons. Also, some persons may have limited dexterity, making use of a stylus, keyboard, or other manual entry mechanism very difficult. The speech

- 10 -

recognition module allows single handed, or even hands free, use since the stylus need not be grasped.

In other embodiments, a voice reception module comprising a microphone and a transmitter can be mounted near the person's mouth, such as on a collar, or the skin, and transmit audio signals to a computing device, such as a PDA, desk-top computer system, remote server, interactive television set top box, and the like. The person can view a menu presented on the display of a computing device, and select items using speech commands directed to the microphone. The voice reception module can further comprise an oscillator, such that the signal received by the microphone is modified by the shape of the person's mouth, for example by the action of forming words (such as letters, numbers, food names) so that, for example, the person can provide commands to the computing device by silently forming words without the need to speak audibly.

#### Diet Log

As described above in relation to Figures 4-6, speech control of menu selections is useful in creating a diet log. In U.S. Pat. Nos. 5,704,350 and 4,891,756, incorporated herein by reference, Williams describes the selection of food items through a hierarchical menu system. This hierarchical approach can be advantageously combined with speech recognition for diet log creation. A diet log software application program was also described in U.S. Provisional Pat. App. No. 60/240,185, filed 10/13/00, incorporated herein by reference. This program can be advantageously modified for use with embodiments of the present invention by adapting the menu selection methods provided by the software program to be responsive to spoken commands, as interpreted by a speech recognition system. Other diet log software are known in the art, for example as described by Kretsch in U.S. Pat. No. 5,233,520.

In embodiments of the present invention, the person is initially presented with a high-level menu of food types, such as a menu listing as shown in Figure 7. As described above, the person can speak the number associated with each food type identifier. The person may also speak a letter corresponding to the first letter of one or more of the menu listings. The menu listing can be reduced to a unique item, which is then selected, or to a number of items each beginning with the spoken letter. The

- 11 -

person can then speak a second letter to further restrict the selection, and a third letter if necessary.

If multiple word listings are presented to the person, a second (or later) spoken letter can be applied to restrict choices to those having a second word beginning with that letter. For example, if breads are listed in a menu as bread, rye; bread, brown; bread, pumpernickel (and the like), a person may speak "B ... R" to select bread, rye.

With respect to Figure 7, a person may speak "1", or "C" to choose cereals. In other embodiments, the vocabulary of the speech recognition system is automatically restricted to the words and numbers displayed on the menu. In this case, the word "cereals" can be reliably distinguished from the other word options by a speech recognition system. In this case, the first syllable alone can be used to distinguish the word "cereals" from the other choices.

Figure 8 shows another menu of food item identifiers which may be presented to the person after the person has chosen cereals from the menu of Figure 7.

Figure 9 shows a portion size menu which may be presented to the person after they have selected a food item identifier.

Figure 10A shows a schematic of a portable computing device which may be used in embodiments of the present invention, comprising microprocessor (processor) 32, memory 34 (which may comprise ROM, RAM, memory modules, and other memory types and combinations), user interface 36 (which may comprise buttons, stylus, and other mechanical data entry mechanisms), a display 38 (comprising a display driver), peripherals 39 (which may comprise an additional display, indicator, printer, buzzer, speaker, modem, drive, and the like), accessory interface 42, and hardware module 44. Figure 10B shows a possible schematic of a hardware module 44 for use in embodiments of the present invention comprising a microphone 46, analog to digital converter 48, signal processor 50, speech analysis circuit 52, and nutritional database 58. In this example, commands such as numeric speech are interpreted by the analysis circuit 52, which sends a corresponding electrical signal to the interface 42, for interpretation by the microprocessor 32. The nutritional database can also be accessed by a software application program running on the portable computing device. In other embodiments, the module 44 may further comprise a processor and additional memory, and diet log software may run on the module,

displaying information on the display 38, and storing a diet log either on the memory of the module, in the memory 34, or at a remote location such as a remote computer accessed over a communications network.

Figure 11 shows a method of creating part of a diet log according to the present invention. Box 70 corresponds to the selection of the operating mode of the computing device, for example by speaking the word "diet". Other words which may be recognized at this stage might include exercise (for exercise logging), balance (for calorie balance calculation), address (for address book), and the like. Box 72 corresponds to the presentation of a diet log welcome screen to the person, possibly including a summary of data entry methods. Box 74 corresponds to the use of available information to modify presentation of menus to the person. Box 76 corresponds to the presentation of types of food to the person, for example as appropriate for the time of day, or other information determined in step 74. Box 78 corresponds to the selection of a type of foods, for example using voice recognition methods. Box 80 corresponds to the display of food items according to the selected food type. Box 82 corresponds to the selection of a food item by identifier, for example by a particular name or code. Box 84 corresponds to the display of a menu of related information. Box 86 corresponds to the selection of related information.

With regard to Box 74, available information can include location, purchase information, time, known preferences, known meals supplied by a weight control business, dietary goals, and the like. The information can be in the memory of the computing device, received over a communication network, preprogrammed, based on previous behavior, or determined using other methods and sources. For example, at times corresponding to lunch times, a lunch menu can be presented. Other meal and snack times can be correlated with the consumption of certain foods, possibly using previously collected data. If the location of the person is known (for example using global positioning, cell phone triangulation, or local wireless network based methods) the menu can be presented accordingly. A person in a particular restaurant can be presented with diet log options corresponding to that restaurant. A person at home can be presented with lists of food known to be at home, for example using prior purchase information or known preferences. If the person is at a business office, typical foods consumed there can be presented as a menu. There is often a strong correlation

- 13 -

between a person's location and the foods that they consume. This correlation can be established by a software learning process, or can be preprogrammed, or entered by the person. The food-location correlations can then be used to restrict menu presentation to the person. If food has been obtained from a vending machine or similar food dispensing machine, or such a machine is detected nearby, menus corresponding to available foods can be presented. Personal expenditure information can also be used, for example if a transaction has taken place at a restaurant or food retail establishment, that information can be used to help determine food consumed. If the identity of a previously purchased meal is known, for example using a credit card transaction log, the identity of the food can be entered into the diet log and the methods of the present invention can then be used to record the amount of food eaten, for example as a fraction of serving size.

With further regard to Box 86, further information can comprise portion size, preparation method, accompanying items or omissions, fraction of meal eaten, and the like. For example, if the person chooses cornflakes as the food item identifier, the person can then be asked to enter portion size and amount of milk. This can be done by presenting a menu of options and receiving spoken choices, possibly using more than one menu depending on the food item selected in box 82.

Figure 12A shows a display (100) of a computing device having a window 102 displaying received speech commands. In this example, the letters A-Z are used for speech commands, but other limited vocabularies such as numbers can also be used. The person has spoken the letter "C", which is recognized by the speech recognition system and displayed at location 104 in window 102. A menu of food item identifiers (or types, according to the diet log method and hierarchy in use) beginning with the letter "C" is displayed. If the person has just eaten cabbage, the person then speaks the letter "A" to distinguish it from the other displayed option. Further letters can be entered to further restrict the menu if necessary, with an item being selected once it is the only menu option remaining. The word "Back" (or similar) can be recognized to delete a displayed letter and remove the corresponding restricted selection of food item identifiers.

Figure 12B shows the letter "K" displayed in window 102, and two diet log options presented, kumquats and pickled kumquats. In this case, the pickled option

- 14 -

can be selected by speaking the letter "P", the first letter of the second word. This general approach is valuable in diet logging, as often a single food item name will be repeated with several modifiers, or plural modifiers. In this case, speaking the first letter of the base name (e.g. "M" for milk), and the first letter of the modifying term(s) (e.g. "S" for skim), can be used for rapid speech-recognition selection of diet log items. The first letter of the base name can be repeated to select the base (unmodified) items. Other letters or numbers can be entered to differentiate ambiguities if necessary. Alternatively, the modifiers can be numbered and the person speaks a number corresponding to the modifier. In more advanced systems, the recognized vocabulary of the speech recognition system is restricted to the options presented to the person, which considerably increases the reliability of the speech recognition system over one trying to match spoken commands to a large available vocabulary. In some embodiments, only the first spoken syllable, or first two syllables, are used for speech recognition.

#### 15 Nutritional Database

The hardware module described above can comprise a memory containing a nutritional database. This database comprises food item identifiers, such as the names of foods, meals, beverages, diet supplements, and the like. The data can be arranged in a hierarchical fashion, so that food item identifiers are grouped by generic type, such as: meat products, dairy products, starches, preferred meals, lunch items, breakfast items, dinner items, meal supplements, nutraceuticals, prepackaged foods, and the like. The generic types can further be grouped into higher level groups, or divided into sub-groups, if convenient. Presented menus can in certain cases comprise both food types and specific food identifiers.

25 A nutrition database can also be contained in the memory of a portable computing device, or on a memory module which may be inserted into the portable computing device or the hardware module described above, using a memory module slot.

The nutritional database can also be located on another computer system, memory module, or device in communication with the portable computing device, as will be described in more detail below.

- 15 -

Food items can be added to the nutritional database using the available data entry mechanisms of the portable computing device, or using data received over a communications network or memory module transfer.

#### Communications Network

5           The portable computing device can be in communication with a remote computer system over a communications network such as the Internet. For example, the portable computing device can be in wireless communication with the Internet through an Internet service provider (ISP), and then through the Internet to a remote server system. (The ISP is considered part of the communications network). In this  
10 case, the nutritional database and speech recognition system can be located on the remote computer system.

          In other embodiments, a voice signal, for example as provided by a microphone or other speech receiving device, can be transmitted over the communications network to a remote computer system, which can comprise a speech  
15 recognition system. A diet log software application can be executed by the portable computing device or remote computer system. Menu displays, nutritional information, diet advice, and the like can be presented on the display of the portable computing device. The necessary bandwidth for uploading voice signals will likely be much less than that necessary for downloading visual data from the remote computer system to  
20 the portable computing device. Hence, different communications networks can be used for data transfer to and from the portable computing device. For example, a wireless phone system can be used to transmit voice data to the remote computer, and a higher bandwidth wireless connection to an Internet service provider can be used to receive data for presentation on the display.

25           Figure 13 shows a portable computing device 120 in communication with a remote server system 124 over communications network 122, which can be the Internet. Spoken commands received by the device 120 are transmitted as digitized audio files to the server 124, which directs them to the speech recognition system 126. The spoken commands are then converted into computer-readable commands  
30 (electronic commands, or speech command signals), and a server software application program can respond to the electronic commands by selecting data from the



- 16 -

nutritional database. Menus generated by the server software application program can be viewed on the display of the computing device 120.

#### Product Codes

- As part of a weight control program, a weight control business can supply a person with prepackaged meals. These can be identified using abbreviated codes for convenient entry into a diet log. For example, a lunch can be identified by a code such as L1, which can be readily entered into a diet log by speech command or other method. The nutritional database can further be supplied by the weight control business, so as to contain nutritional information correlated with each product code.
- Product identities and/or nutritional content can be algorithmically converted to a numeric (or alphanumeric) code, analogous to Gemstar VCR Plus codes. This facilitates the entering of this simplified data into a diet log by speaking the code characters.

#### Exercise Log

- The person can also record exercises on the portable computing device in an analogous manner. Exercises and activities can be grouped in a hierarchical fashion within an exercise database, contained within the memory of the portable computing device, a module, or other device in communication with the portable computing device. The portable computing device can have several operating modes, such as diet log, exercise log, calorie balance calculator, time display, position display, organizer, word processor, and the like, and can be switched between modes by any convenient method.

#### Calorie Balance

- The person can determine their resting metabolic rate (RMR) using a metabolic rate meter such as an indirect calorimeter. The RMR can be entered into the portable computing device by any convenient method, such as speaking the digits combined with speech recognition. The calorie management software can then be used to calculate calorie balance using diet log entries, exercise log entries (and/or estimated activity levels), and RMR.

### PDA-Cell Phone

As is known in the art, a portable computing device such as a PDA can be readily combined with the functionality of a wireless telephone into a unitary device, which we will term a DAP (digital assistant - phone). In this case, a microphone will  
5 be present in the DAP and can be used for speech command operation of a diet log software. The speaker of the DAP can be used to provide feedback to the person, such as noises to indicate successful or unsuccessful speech recognitions. The term portable computing device, as used herein, includes devices with additional functionality such as a DAP.

10 Using a portable computing device with a speaker, speech synthesis can be used to provide a dialog with the person, so that a display may not be needed for menu presentation. In this case, menus can be presented using synthesized speech. For example, a person speaks the word "Diet" to the portable computing device. The device responds "Diet Log Activated. It is now 12:10 pm. Lunch food types include 1  
15 - burrito, 2 - sandwich, 3 - yogurt, 4 - other. Please speak number". The person speaks the number "1". The device responds "Burrito chosen. Choose 1 - bean, 2 - beef, 3 - chicken.". The person speaks the number "3". The device responds "Did you eat the whole chicken burrito". The person responds "Yes". The device responds "OK. Chicken burrito entered into diet log. Don't forget your run this afternoon."

### 20 Interactive Television

A portable device can also be used for remote control of a digital interactive television or other entertainment device. Spoken commands can be transmitted to the set-top box of the digital interactive television, for example by speaking to the portable device with wireless transmission of a digitized audio file to the set-top box.  
25 The set-top box, or other device such as a remote server in communication with it, has a voice recognition system to convert the digitized audio file to electronic data, such as text and/or numbers. The electronic data can be used for channel selection of the digital interactive television, or selection from menus such as food item menus displayed on the digital interactive television. Spoken commands can be transmitted  
30 from a portable device, possibly with the form factor of a remote control, over a wireless telephone network to a remote computer having a speech recognition system,

- 18 -

which then responds to the received and interpreted commands by providing corresponding visual menu presentations, selection indications, channel selections, program selections, and the like on the display of the interactive television.

#### Physical Location Determination

5           The portable computing device, or add-on hardware module, can further comprise a physical location determining unit, such as a global positioning system, triangulation system, or the like. The location can be correlated with positions of food retail establishments using a food retail database, and so with menus and available food lists. Hence, a portable computing device according to the present invention can  
10       present a food menu to a person based on the person's location. If the person is in a restaurant, a menu can be presented on a display of the portable computing device. The person can then select from the presented menu using the speech recognition systems described herein. A nutritional database is then used to assist in creating a diet log from the selected items.

#### 15       Wrist-Mounted Device

          Embodiments of the present invention can comprise a portable device supported by a strap around a wrist of the person, in the form of a wristwatch. The display of the device can be used to display time, other data, entertainment, and food menus to the person. A microphone in the housing can be used to receive spoken  
20       commands. A speech signal, such as a digitized audio file, can be transmitted to other devices carried by the person, or with which the person interacts (for example a PDA, desktop computer, entertainment device, remote computer, and the like). The display of menu options can be presented on the display of another device, such as a PDA.

#### Limited Display Capabilities

25           In some cases, display complexity will be insufficient to display all options to the person in the form of a menu. In this case, the person can speak the first letter of the required option, so as to limit the display to those food items or food types beginning with that letter. Other letters or numbers can then be spoken so as to further limit the display, or to select an item.

- 19 -

Other variations on the disclosed embodiments of the present invention will be apparent to those of skill in the art. These variations do not depart from either the spirit or scope of the present invention, and therefore the disclosed embodiment should be interpreted broadly.

- 20 -

5           Having described my invention, I claim:

2           1.     A speech recognition module for a portable computing device,  
2     comprising:

4           a module housing configured to engage with an accessory interface of the  
4     portable computing device;

6           a speech receiving device, adapted to receive a speech command from the  
6     person and to provide an electrical speech signal; and

8           a speech recognition system, adapted to receive the electrical speech signal  
8     from the speech receiving device, to distinguish the speech command, and to  
10    communicate a corresponding electrical speech command signal to the portable  
10    computing device, wherein the electrical speech command signal is used to control  
12    the operation of a software application program running on the portable computing  
12    device.

2           2.     The speech recognition module of claim 1, further comprising a  
2     nutritional database, the nutritional database comprising food item identifiers and  
associated nutritional data for a plurality of food items.

2           3.     The speech recognition module of claim 2, wherein the associated  
2     nutritional data comprises a calorie content associated with each of the food item  
identifiers.

2           4.     The speech recognition module of claim 2, wherein the software  
2     application program is adapted to select a selected food item identifier from the  
nutritional database under control of the electrical speech command signal.

2           5.     The speech recognition module of claim 4, wherein the software  
2     application program is further adapted to store the selected food item identifier and  
associated nutritional data in a memory so as to create a diet log for the person.

2           6.     A portable computing system, comprising:  
2     a display;  
a microprocessor;

- 21 -

4 a memory;  
a database of nutritional data relating to each of a plurality of food item  
6 identifiers;  
a speech recognition system providing an electrical speech command signal  
8 correlated with a spoken input;  
a software application program, executed by the microprocessor, adapted to  
10 display a food menu of food item identifiers on the display, to select a food item  
identifier from the food menu based on the electrical speech command signal received  
12 from the speech recognition system, and to store nutritional data relating to the  
selected food item identifier in the memory so as to create a diet log of selected food  
14 items.

7. The portable computing device of claim 6, wherein the database of  
2 nutritional data relating to each of a plurality of food item identifiers comprises a  
caloric content of food items represented by the food item identifiers.

8. The portable computing device of claim 6, wherein the database is  
2 arranged in a hierarchical form wherein particular food items are characterized by a  
generic type, and wherein the software application program is further adapted to  
4 display a menu of generic types on the display, and to select a generic type from the  
menu of generic types on receipt of the electrical speech signal.

9. An accessory module for a portable computing device, the portable  
2 computing device comprising a microprocessor, a display, and an accessory port, the  
accessory module comprising:  
4 an electrical interface, adapted to communicate with the accessory port;  
a memory;  
6 a nutritional database, stored in the memory, correlating each of a plurality of  
food item identifiers with an associated food item nutritional content;  
8 a software application program, stored in the memory, and executable by the  
microprocessor so as to present a menu of food item identifiers on the display, further  
10 adapted to select a food item selection from the menu of food item identifiers on

- 22 -

receipt of a speech command signal, and further adapted to store the food item  
12 selection and correlated nutritional content within a diet log for the person.

10. The accessory module of claim 9, wherein the associated food item  
2 nutritional content for each of the plurality of food item identifiers comprises a calorie  
content for each of the plurality of food item identifiers.

11. The accessory module of claim 9, wherein the speech command signal  
2 is provided by a speech recognition system resident on the accessory module.

12. The accessory module of claim 9, wherein the speech command signal  
2 is provided by a speech recognition system resident on the portable computing device.

13. A method of creating a diet log for a person, comprising:  
2 presenting a menu of food type identifiers to the person;  
receiving a type selection speech command from the person;  
4 correlating the type selection speech command with a selected food type  
identifier;  
6 presenting a menu of food item identifiers to the person, wherein the food item  
identifiers have a hierarchical relationship to the selected food type identifier;  
8 receiving an item selection speech command from the person;  
correlating the item selection speech command with a selected food item  
10 identifier; and  
storing the selected food item identifier and associated nutritional data within  
12 a memory device, so as to create a diet log for the person.

14. The method of claim 13, wherein the selected food item identifier is  
2 correlated with associated nutritional data using a nutritional database.

15. The method of claim 13, wherein the menus are presented to the  
2 person on the display of a portable computing device carried by the person.

- 23 -

16. The method of claim 15, wherein the menus presented to the person  
2 are chosen based on the physical location of the person.

17. The method of claim 15, wherein the correlation of speech commands  
2 with selected identifiers is performed by a software application program running on  
the portable computing device.

18. The method of claim 15, wherein the correlation of speech commands  
2 with selected identifiers is performed by a software application program running on a  
remote computer system in communication with the portable computing device.



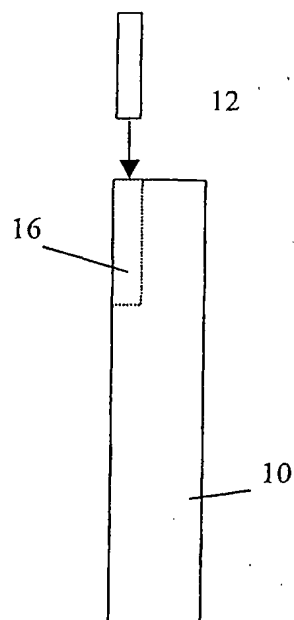


Figure 1

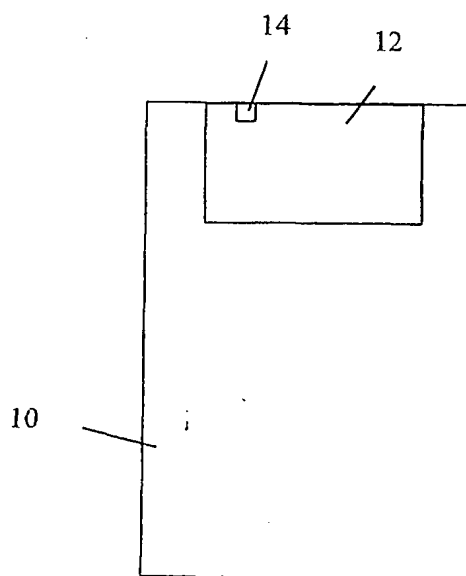


Figure 2

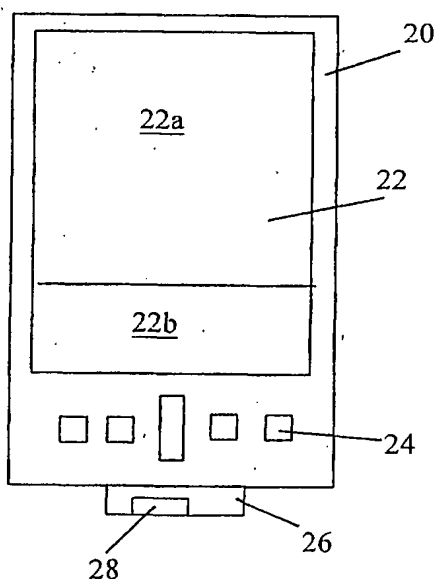


Figure 3

<b>Meal Type</b> 1) Breakfast 2) Lunch 3) Dinner 4) Snack 5) Drink 6) Custom
Speak or Press Number
<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>

Figure 4

<b>Breakfast</b> 1) Cold Cereal 2) Hot Cereal 3) Egg Based meal 4) Pancakes/Waffles 5) Other
Speak or Press Number
<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>

Figure 5

<b>Hot Cereal</b> 1) Oatmeal 2) Malt-o-Meal 3) Grits 4) Other
Speak or Press Number
<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>

Figure 6

<b>Food Type</b> 1) Cereals 2) Dairy products 3) Meats 4) Breads, 5) Vegetables 6) Fruit
Speak or Press Number
<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>

Figure 7

<b>Cereal Name</b> 1) Branflakes 2) Cornflakes 3) Muesli 4) Frosted Flakes 5) Sugar Bombs
Speak or Press Number
<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>

Figure 8

<b>Portion size</b> 1) 0.5 cup 2) 0.75 cup 3) One cup
Speak or Press Number
<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>

Figure 9

3 / 5

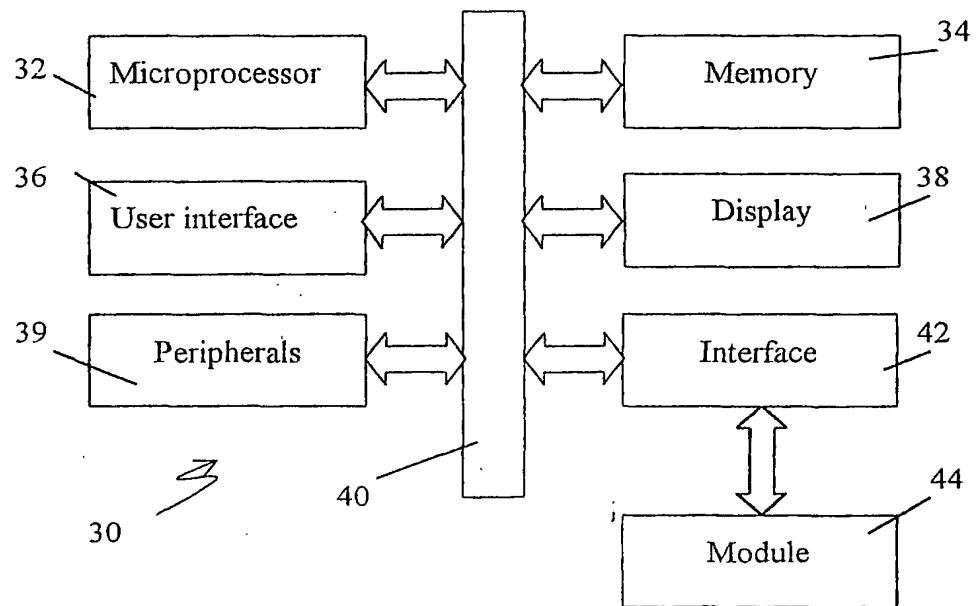


Figure 10A

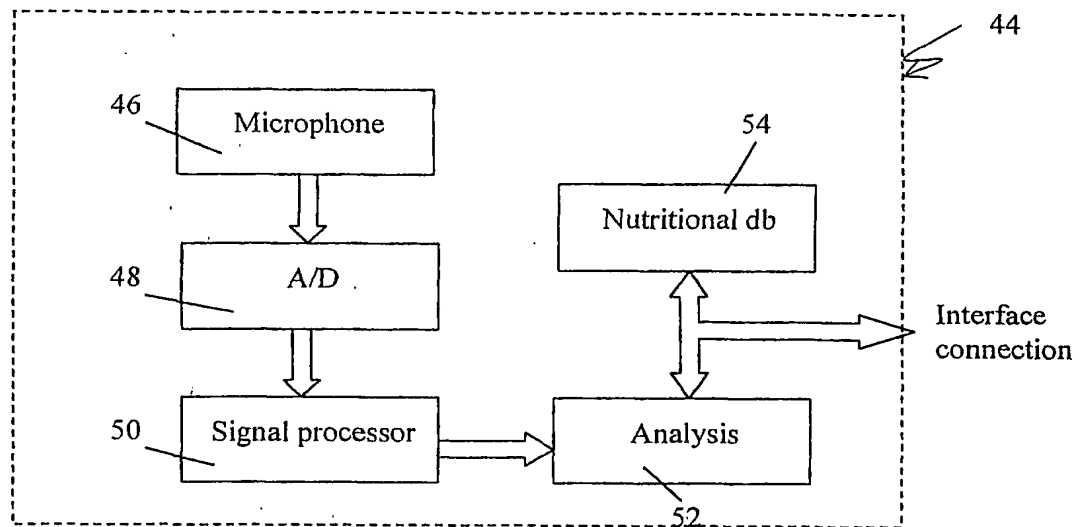


Figure 10B

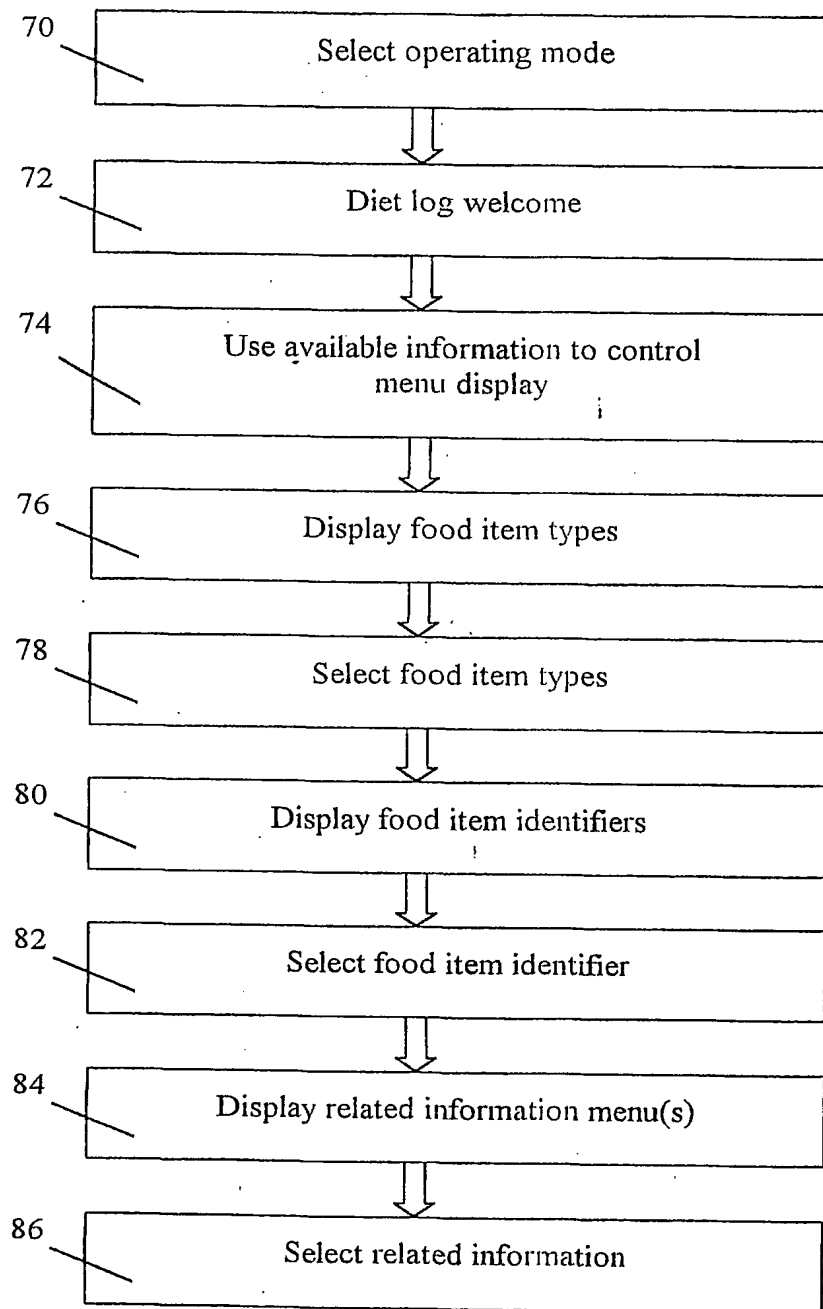


Figure 11

5 / 5

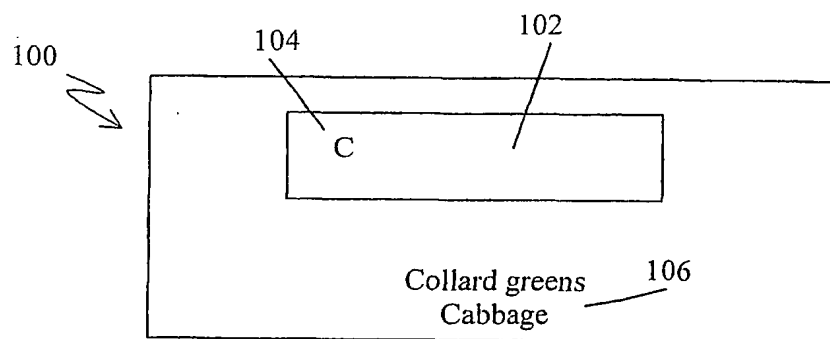


Figure 12A

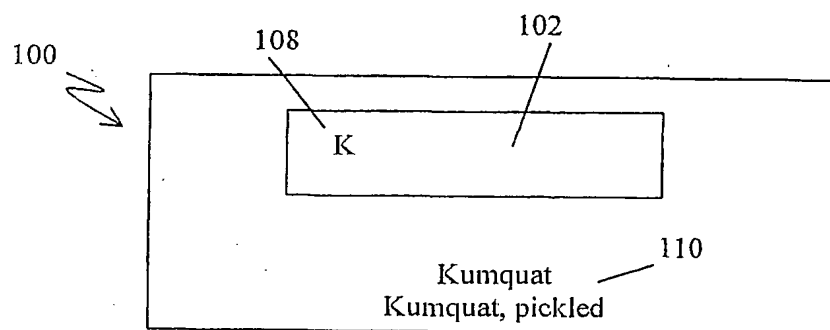


Figure 12B

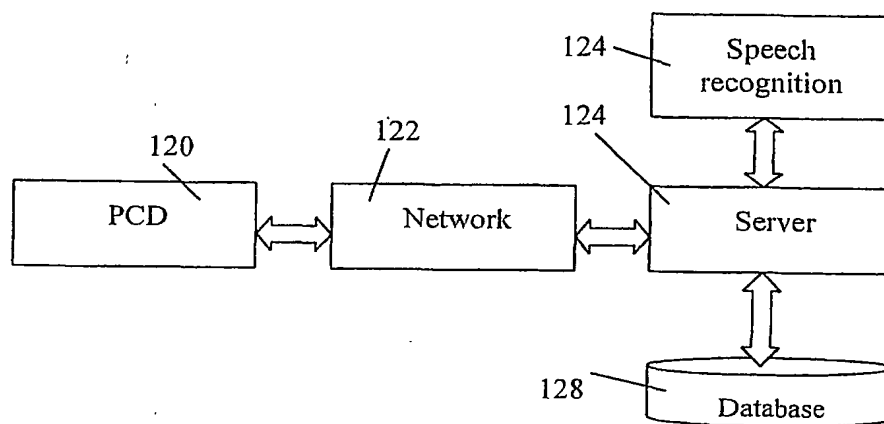


Figure 13

(19) World Intellectual Property Organization  
International Bureau



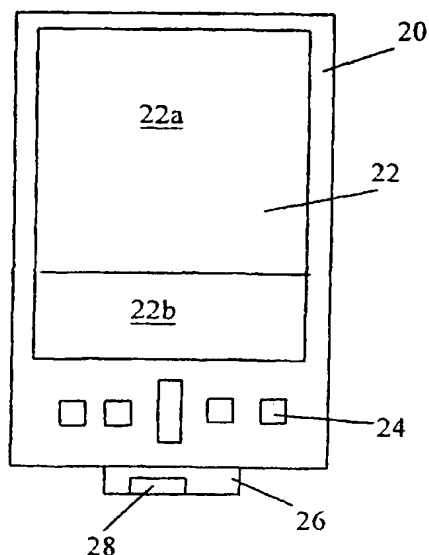
(43) International Publication Date  
20 December 2001 (20.12.2001)

PCT

(10) International Publication Number  
**WO 01/97211 A3**

- (51) International Patent Classification<sup>7</sup>: **G10L 21/00**
- (21) International Application Number: **PCT/US01/41031**
- (22) International Filing Date: **18 June 2001 (18.06.2001)**
- (25) Filing Language: **English**
- (26) Publication Language: **English**
- (30) Priority Data:  
**60/212,319** **16 June 2000 (16.06.2000)** **US**
- (71) Applicant (*for all designated States except US*): **HEALTHTECH, INC.** [US/US]: 523 Park Point Dr., 3rd Floor, Golden, CO 80401 (US).
- (72) Inventor; and
- (75) Inventor/Applicant (*for US only*): **MAULT, James, R.** [US/US]: 1580 Blakecomb Court, Evergreen, CO 80439 (US).
- (74) Agents: **KRASS, Allen, M.** et al.; Gifford, Krass, Groh, Sprinkle, Anderson & Citkowski, PC, Suite 400, 280 N. Old Woodward Ave., Birmingham, MI 48009 (US).
- (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (*regional*): ARIPO patent (GI, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
- Published:**  
— *with international search report*
- (88) Date of publication of the international search report:  
**20 June 2002**
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: **SPEECH RECOGNITION CAPABILITY FOR A PERSONAL DIGITAL ASSISTANT**



(57) Abstract: A speech recognition module for a personal digital assistant comprises: a module housing designed to engage with an accessory feature of the PDA, such as an accessory slot; a microphone (28) for receiving speech commands from the person; and a speech recognition system. A corresponding electrical speech command signal is communicated to the portable computing device (20), allowing control of the operation of a software application program running on the portable computing device. In particular, menu items can be selected for generation of a diet log for the person, for example during a weight control program (22).

WO 01/97211 A3

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US01/41031

**A. CLASSIFICATION OF SUBJECT MATTER**IPC(7) : G10L 21/00  
US CL : 704/275

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 704/275, 270, 270.1, 276, 252, 231

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched  
NoneElectronic data base consulted during the international search (name of data base and, where practicable, search terms used)  
WEST, PLUS, Smart patent workbench**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 4,545,023 A (MIZZI) 01 October 1985 (01.10.1985), Col.1, lines 23-60.	1, 6, 9, 13
Y	US 5,524,169 A (COHEN et al.) 04 June 1996 (04.06.1996), abstract.	2-5, 7-8, 10-12, 14-18
Y	US 5,749,072 A (MAZURKIEWICZ et al.) 05 May 1998 (05.05.1998), abstract, Fig.4.	1-18



Further documents are listed in the continuation of Box C.



See patent family annex.

Special categories of cited documents:	
* "A" document defining the general state of the art which is not considered to be of particular relevance	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
* "E" earlier application or patent published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
* "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
* "O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family
* "P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

21 August 2001 (21.08.2001)

Date of mailing of the international search report

02 JAN 2002

Name and mailing address of the ISA/US

Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703)305-3230

Authorized officer

William R. Korzuch

Telephone No. (703) 305-4700

Form PCT/ISA/210 (second sheet) (July 1998)